Communication Complexity and Energy Efficient Consensus Algorithm

Yilin Mo, Bruno Sinopoli

Dept of Electrical and Computer Engineering, Carnegie Mellon University

Introduction

- Average consensus problems have been extensively studied by many researchers over the past few years.
- Usually consensus algorithms are designed to achieve the fastest convergence rate per iteration. Is it a good metric?
- The time needed for one iteration varies depending on the particular algorithm used.
- For certain application, e.g. wireless sensor networks, energy constraints may be more important than real-time requirements.
- Need to develop a new metric to assess energy efficiency of a consensus algorithm.

Deterministic and Gossip Algorithm

- Consensus algorithms are usually classified into two categories: deterministic or stochastic.
- In this paper we characterize the energy efficiency of deterministic and gossip algorithms.
- We model the network as a connected undirected graph $G = \{V, E\}$.
- Deterministic Algorithm:
  - Update Equation:

$$x_{k+1} = Px_k. \tag{1}$$

  - If P satisfies the following conditions, then average consensus will be achieved:
    1. $\lambda_1(P) = 1$ and $|\lambda_i(P)| < 1$ for all $i = 2, \dots, N$.
    2. $P\mathbf{1} = \mathbf{1}$, i.e. $\mathbf{1}$ is an eigenvector of P.
  - Moreover we assume P is symmetric and non-negative.
  - The average number of communications per node for each iteration is defined as:

$$\bar{d}(P) \triangleq \sum_{i\neq j} \mathbb{I}_{\{P_{ij}\neq 0\}}/N. \tag{2}$$

- Gossip Algorithm:
  - For each iteration, a pair of nodes (i, j) is selected with probability Q_{ij}.
  - The pair exchanges information and updates its states to be the average of the two.
  - Define

$$W_{ij} = I - (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)'/2. \tag{3}$$

  where $\mathbf{e}_i \in \mathbb{R}^N$ is a vectors of all zeros with only the ith element equal to 1.
  - The update equation:

$$x_{k+1} = W_k x_k, \tag{4}$$

  where W_k is a random matrix and the probability that W_k equals W_{ij} is P_{ij}.
  - We assume Q satisfies the following properties:
    1. $\mathbf{1}'Q\mathbf{1} = 1$.
    2. Q is symmetric and non-negative.
- The accuracy of consensus at kth step is defined as:

$$\varepsilon_k \triangleq \sup_{y_0\neq 0} y_k'y_k/(y_0'y_0). \tag{5}$$

Communication Complexity

- Communication complexity measures the average number of communications needed to reach a specified accuracy.
- Let the accuracy be $\varepsilon > 0$. Stopping time T_ε is defined as

$$T_\varepsilon \triangleq \inf\{k : \mathbb{E}\varepsilon_k \leq \varepsilon\}. \tag{6}$$

- Define c_k to be the number of communications incurred at the k^{th} iteration.
- We will indicate with Ω and ω the complexities of deterministic and gossip algorithms respectively.
- We define communication complexity as:

$$\Omega = \omega = \limsup_{\varepsilon\to 0^+} -\frac{\mathbb{E}\sum_{k=0}^{T_\varepsilon} c_k}{\log(\varepsilon)}. \tag{7}$$

- The goal: Find the consensus algorithm with the lowest communication complexity.

Complexity of Deterministic Algorithms

- For the deterministic consensus, Ω is given by

$$\Omega(P) == -\frac{N\bar{d}(P)}{2\max_{i=2,\dots,N}\log(|\lambda_i(P)|)}. \tag{8}$$

- $\Omega(P)$ is hard to minimize since it is in fractional form and contains $\bar{d}(P)$.

Complexity of Gossip Algorithms

- We define the projection matrix $\mathcal{P}$ as

$$\mathcal{P} \triangleq I - \mathbf{1}\mathbf{1}'/N, \tag{9}$$

- the matrix $\mathcal{W}_{ij}$ as

$$\mathcal{W}_{ij} \triangleq \mathcal{P}W_{ij}\mathcal{P} \tag{10}$$

- and the linear operator $\mathcal{A}_Q$ from $\mathbb{R}^{N\times N}$ to $\mathbb{R}^{N\times N}$ as

$$\mathcal{A}_Q(X) \triangleq \sum_{i,j} Q_{ij}\mathcal{W}_{ij}X\mathcal{W}_{ij}. \tag{11}$$

  Let us define the spectral radius of the above operator as $\rho(Q)$.
- For the gossip algorithm, ω is given by

$$\omega(Q) = -\frac{2\sum_{i\neq j} Q_{ij}}{\log(\rho(Q))}. \tag{12}$$

- $\omega(Q)$ is still in fractional form. However we can prove the following inequality

$$\omega(Q) \geq \omega(\tilde{Q}), \tag{13}$$

  where $\tilde{Q}$ is defined as

$$\tilde{Q} = \frac{1}{\sum_{i\neq j} Q_{ij}}[Q - diag(Q)]. \tag{14}$$

  Removing the null operation can reduce communication complexity.
- Optimizing $\omega(Q)$ is equivalent to solving the following problem:

$$\begin{array}{ll} \underset{Q\in\mathcal{S}}{\text{minimize}} & \rho(Q) \\ \text{subject to} & \mathbf{1}'Q\mathbf{1} = 1,\ Q_{ii} = 0, \end{array}$$

  which is convex and can be solved efficiently.

Communication Complexity Comparison between Deterministic and Gossip Algorithms

- Finding the optimal gossip algorithm is easy while finding the optimal deterministic algorithm is in general a hard problem.
- Can we compare the energy efficiency of these two algorithms?
- There exists a natural mapping between deterministic and gossip algorithms.

$$\begin{aligned} f : & P \to Q \\ & P \mapsto P/N. \end{aligned}$$

- The following condition is sufficient for $\Omega(P) \geq \omega(P/N)$

$$\lambda_2(P) \geq \frac{16}{\bar{d}(P)^2}. \tag{15}$$

- Inequality (15) is true for a large class of networks. The main reason is that the condition does not depend on the size of the graph N.
- For most graphs, the gossip algorithm is more energy efficient than the deterministic one.

Illustrative Examples

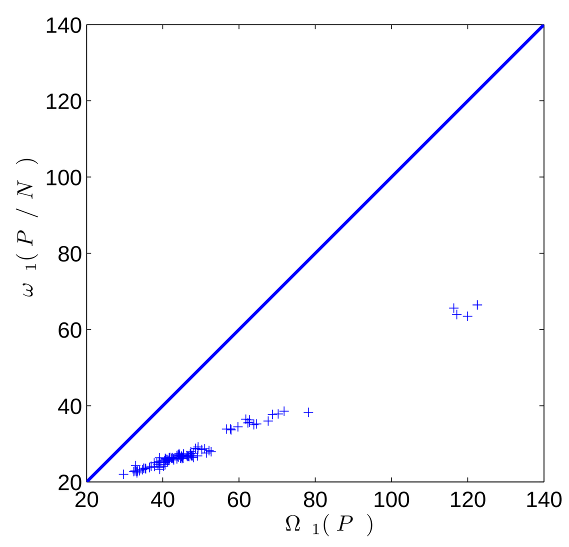

Figure: $\Omega(P)$ v.s. $\omega(P/N)$

- We use 100 randomly generated connected graphs of 10 vertices and 50 edges.
- The consensus matrix P is chosen of the following form:

$$P = I - \alpha L,$$

  where L is the Laplacian matrix of the graph with eigenvalues $\lambda_1(L) \geq \ldots \geq \lambda_{N-1}(L) > \lambda_N(L) = 0$ and

$$\alpha = 2/(\lambda_1(L) + \lambda_{N-1}(L)).$$

Conclusion

- A new energy metric for consensus algorithms is defined and explicit formulas are provided to compute the communication complexity for both deterministic and gossip algorithms.
- Finding the optimal gossip algorithm with minimum communication complexity is formulated as a convex optimization problem. A non convex optimization problem needs to be solved to find its deterministic counterpart.
- A comparison between the complexity of deterministic and gossip algorithms is also provided, showing that gossip-based consensus is more desirable than deterministic consensus if energy efficiency is the main objective.